IN THE CLAIMS:

Kindly replace the claims of record with the following full set of claims:

1. (Currently amended) A method of calculating iteration values for free parameters λ_α in a maximum-entropy speech model in a speech recognition system in accordance with the following general training algorithm:

$$\lambda_\alpha^{(n+1)}\Big|_{\alpha \in A_{i(n)}} = G(\lambda_\alpha^{(n)}, m_\alpha, m_\alpha^{(n)}, \ldots)\Big|_{\alpha \in A_{i(n)}}$$

where:

n: refers to an iteration parameter that represents a current iteration step;

A_i: represents the i-th attributes group in the speech model, where $1 \leq i \leq m$;

$A_{i(n)}$:represents the attributes group selected in the n-th iteration step;

α: represents an attribute in the speech model;

G: represents a mathematical function;

$\lambda_\alpha^{(n)}$: represents the n-th iteration value for the free parameter λ_α ;

m_α : represents a desired boundary value for the attribute α; and

$m_\alpha^{(n)}$: represents the n-th iteration boundary value for the desired boundary value m_α, where one attribute group $A_{i(n)}$ from a total of m speech model attribute groups is assigned to each iteration parameter n, and where the iteration values $\lambda_\alpha^{(n+1)}$ are calculated for each attribute α from the currently assigned attribute group $A_{i(n)}$ characterized in that the current iteration parameter n is assigned the attribute group $A_{i(n)}$ where $1 \leq i(n) \leq m$, for which, in accordance with a predefined criterion, the adaptation of the iteration boundary values $m_\alpha^{(n)}$ to the respective associated desired boundary values m_α, is the worst of all the m attribute groups of the speech model, wherein the predefined criterion is determined before each incrementation of the iteration parameter n by:

BEST AVAILABLE COPY

Amendment After Final Rejection
Serial No. 10/075,866

Docket No. DE010031

a) Calculating current iteration boundary values $m_\alpha^{(n)}$ for attributes α from all the attribute groups Ai, where 1 ≤ i ≤ m, of the speech model according to the following formula:

$$m_\alpha^{(n)} = \sum_{(h,w)} N(h) \cdot p^{(n)}(w|h) \cdot f_\alpha(h,w)$$

where $N(h)$: describes the frequency with which the string of words h (history) occurs in a speech model-training corpus;

p(n) (w | h): is an iteration value for the probability with which the word w follows the history h; and

f_α (h,w): represents an attribute function for the attribute α;

b) Selecting the attribute group $A_{i(n)}$ for which the iteration boundary values $m_\alpha^{(n)}$ are most poorly adapted to the associated boundary values m_α by executing the following steps:

bi) for each attributes group A_i: Calculating the criterion $D_i^{(n)}$ according to the following formula:

$$D_i^{(n)} = \left[\sum_{\alpha\in Ai} t_\alpha m_\alpha \cdot \log\left(\frac{m_\alpha}{m_\alpha^{(n)}}\right) + \left(1 - \sum_{\alpha\in Ai} t_\alpha m_\alpha\right) \cdot \log\left(\frac{1 - \sum_{\alpha\in Ai} t_\alpha m_\alpha)}{1 - \sum_{\alpha\in Ai} t_\alpha m_\alpha^{(n)})}\right)\right];$$

where t_α represents a convergence increment;

bii) Selecting the attribute group $A_{i(n)}$ with the largest value for the criterion $D_i^{(n)}$ according to:

$$i(n) = \arg\max_j D_j^{(n)}$$

biii) Updating the parameter $\lambda_\alpha^{(n+1)}$ for all the attributes α from the selected attribute group $A_{i(n)}$ and

c) Repeating steps a) and b) in each further iteration step, until all boundary values $m_\alpha^{(n+1)}$ converge with a desired convergence accuracy.

2. (Cancelled)

3. (Currently amended) A method as claimed in Claim [[2]] 1, characterized in that the following initialization steps are carried out before the first execution of steps a) through c):

a') Determining values for the convergence increments α; and

a'') Initializing p(0)(w | h) with any set of parameters $\lambda_\alpha^{(0)}$.

4. (Currently amended) A method as claimed in Claim 3, characterized in that the values of the convergence increments [[tα]] t_α for each attributes group A_i are calculated in step a') as follows:

$$t_\alpha = \frac{1}{M_i}$$

with

$$M_i = \max_{(h,w)}\left(\sum_{\alpha \in A_i} f_\alpha(h,w)\right)$$

5. (Previously presented) A method as claimed in claim 1, characterized in that the function G represents a Generalized Iterative Scaling (GIS) training algorithm, and is defined as follows:

$$\lambda_\alpha^{(n+1)} = G = \lambda_\alpha^{(n)} + t_\alpha \cdot \log\left(\frac{m_\alpha}{m_\alpha^{(n)}} \cdot \frac{1-\sum_{\beta \in Ai(n)} t_\beta . m_\beta^{(n)}}{1-\sum_{\beta \in Ai(n)} t_\beta \cdot m_\beta}\right),$$

where

α represents a specific attribute and β all the attributes from the selected attribute group $A_{i(n)}$.

6. (Currently amended) A method as claimed in claims [[2]] 1, characterized in that the attribute function f_α is an orthogonalized attribute function f_α^{ortho}, which is defined as follows:

$$f_\alpha^{ortho}(h,w) = \begin{cases} 1 & \text{if } \alpha \text{ is the attribute with the highest range in Ai which correctly describes the string of words (h,w)} \\ 0 & \text{otherwise} \end{cases}$$

7. (Previously presented) A method as claimed in Claim 6, characterized in that a desired orthogonalized boundary value m_α^{ortho} is calculated according to:

$$m_\alpha^{ortho} = m_\alpha - \sum_{(*)} m_\beta^{ortho}$$

where (*) contains all the higher ranging attributes β which include the attribute α and which come from the same attribute group as α.

8. (Previously presented) A speech recognition system comprising a recognition device for recognizing the semantic content of an acoustic signal, in particular a voice signal, recorded by a microphone and made available, by mapping parts of this signal onto predefined recognition symbols as supplied by the maximum entropy speech model MESM, and for generating output signals which represent the recognized semantic content; and a training arrangement- for adapting the MESM to recurring statistical patterns in the speech of a specific user of the speech recognition system, characterized in that the training arrangement calculates free parameters λ in the MESM in accordance with the method as claimed in Claim 1.

9. (Previously presented) A training arrangement for adapting the maximum entropy speech model (MESM) in a speech recognition system to recurring statistical patterns in the speech of a specific user of the speech recognition system, characterized in that the training arrangement calculates free parameters λ in the MESM in accordance with the method as claimed in Claim 1.